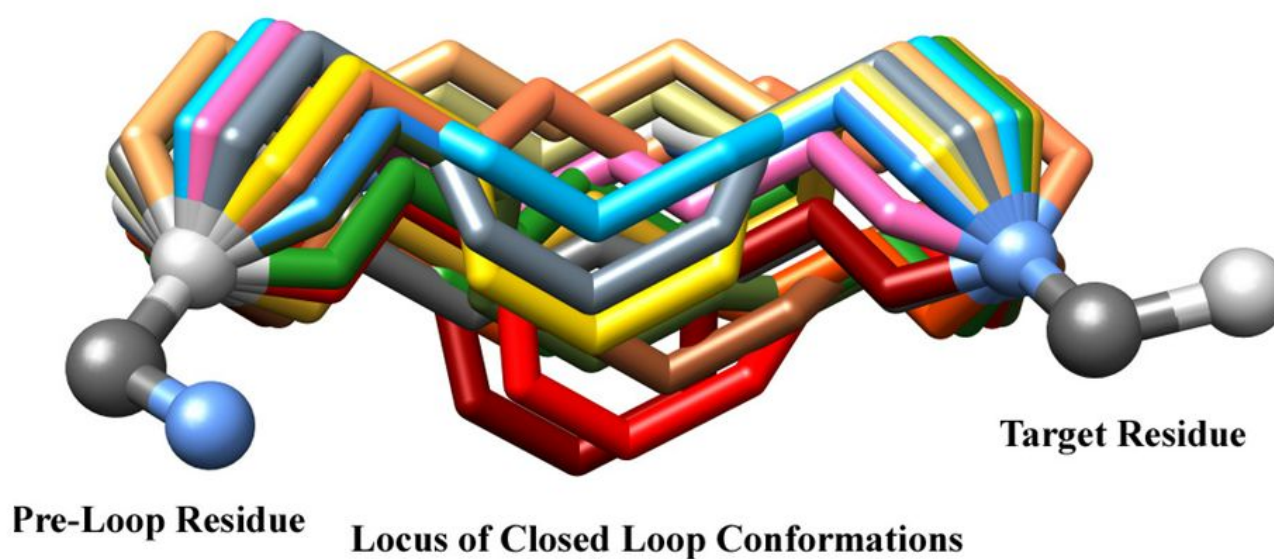
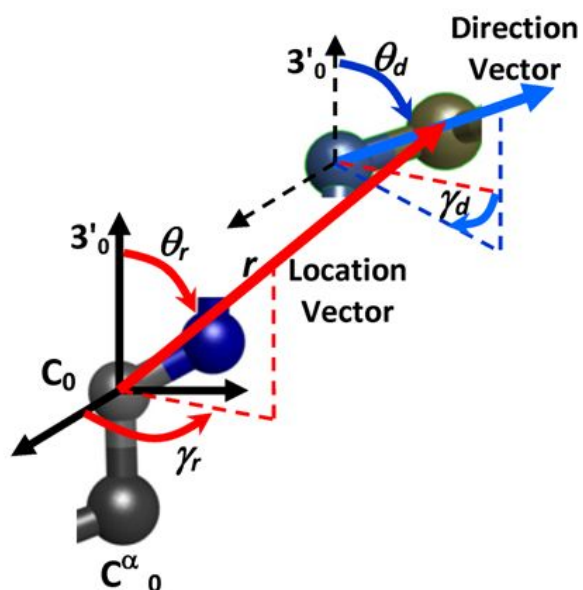


Loop modeling and protein refinement



Agenda

- ✓ Motivations
 - ✓ Loop modeling
 - modeller
 - Fread
 - KIC
 - Sphinx
 - ✓ Protein refinement
 - energy minimisation
-

Chothia and Lesk plot

Sequence vs Structure conservation

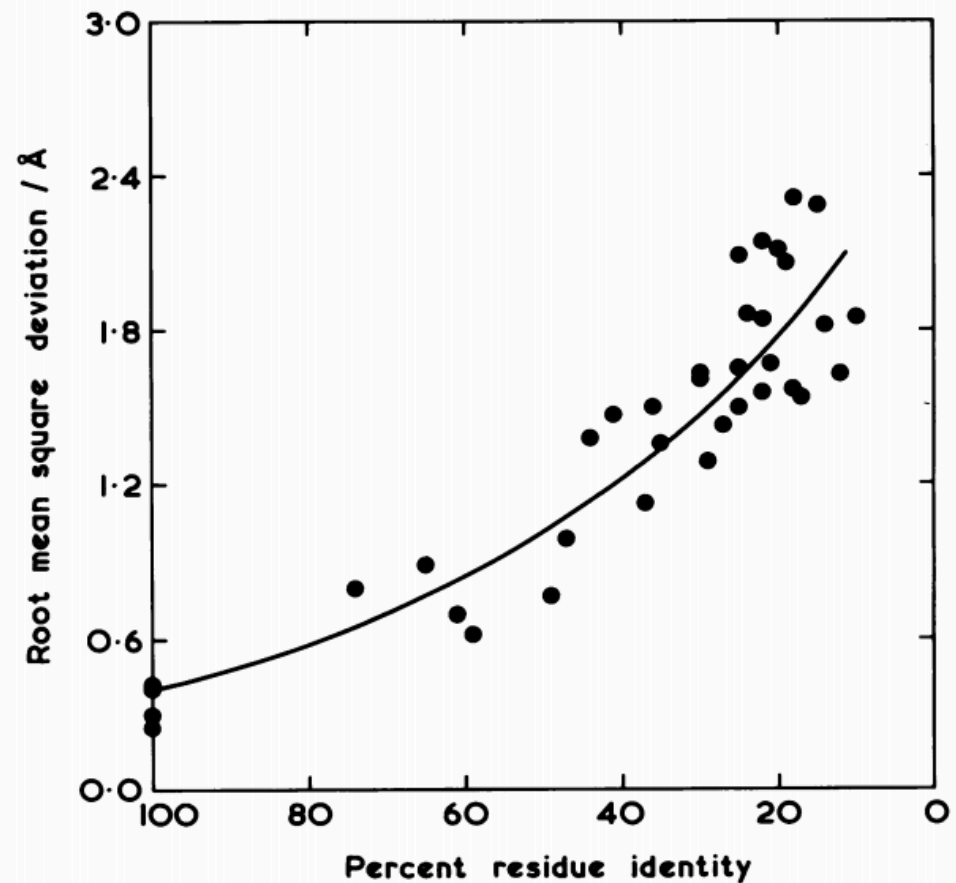


Fig. 2. The relation of residue identity and the r.m.s. deviation of the backbone atoms of the common cores of 32 pairs of homologous proteins (see Table II).

Chothia and Lesk plot

Sequence vs Structure conservation

It is still valid

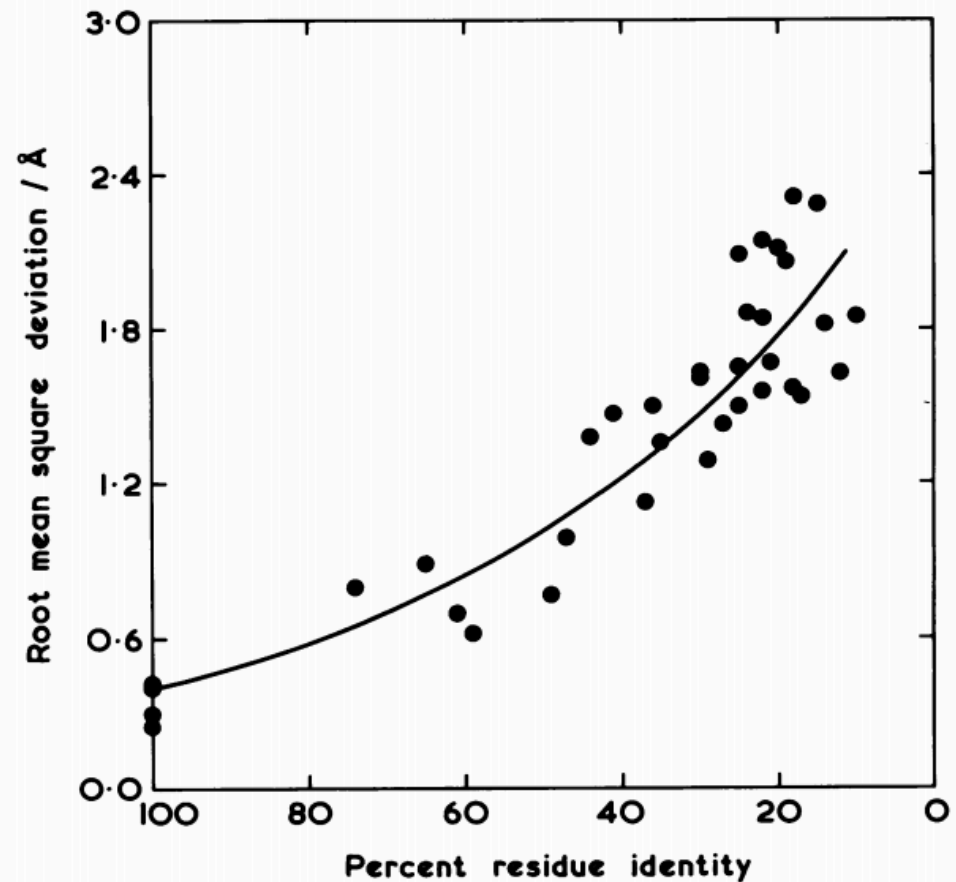


Fig. 2. The relation of residue identity and the r.m.s. deviation of the backbone atoms of the common cores of 32 pairs of homologous proteins (see Table II).

Chothia and Lesk plot

Sequence vs Structure conservation

It is still valid

Not for loops!!!

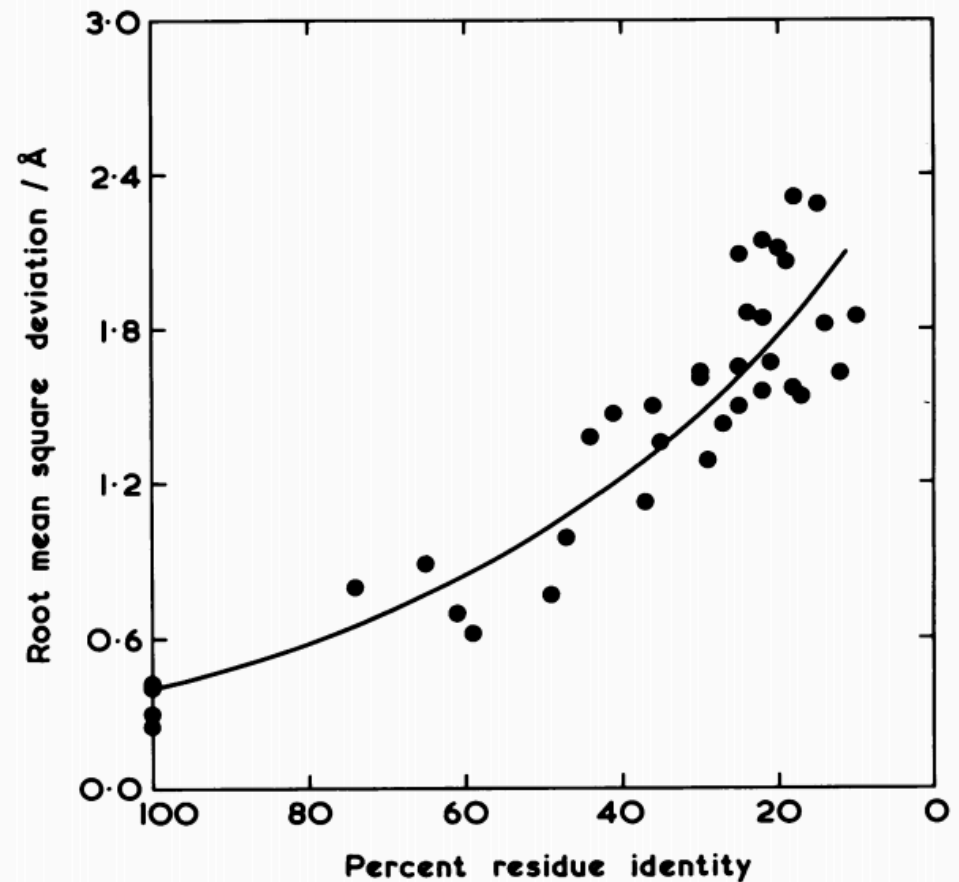


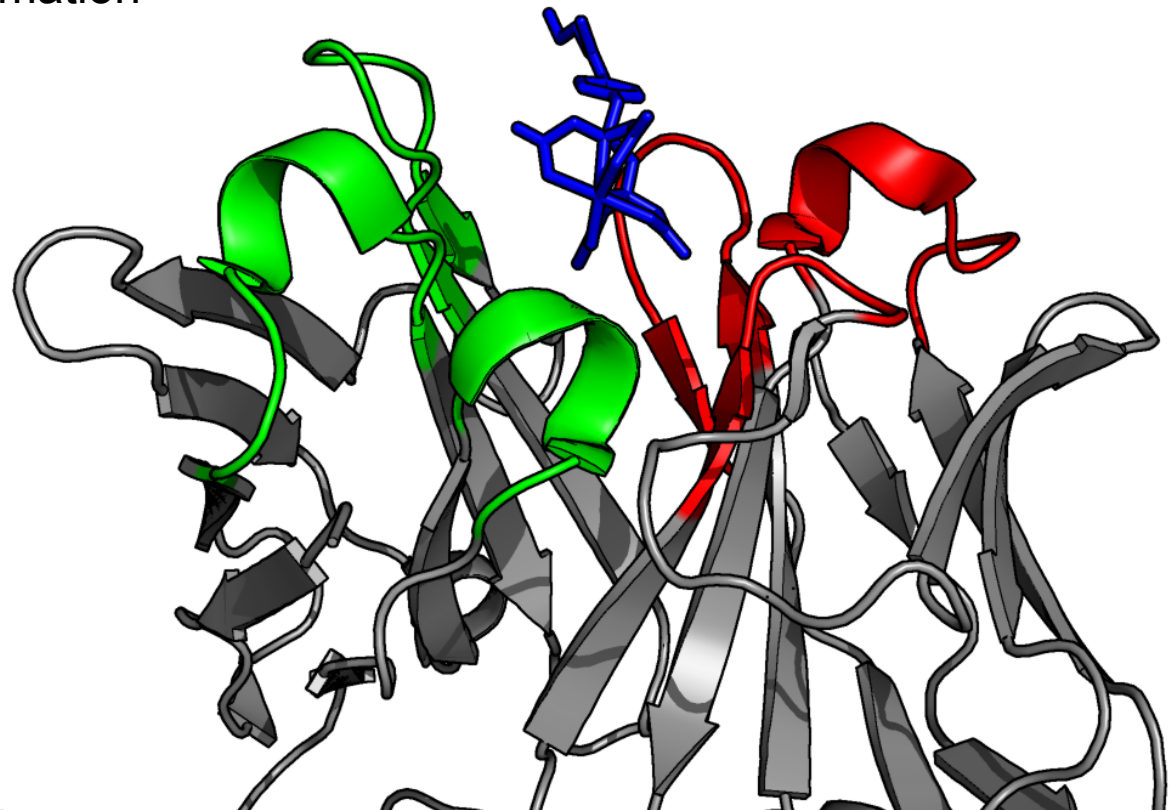
Fig. 2. The relation of residue identity and the r.m.s. deviation of the backbone atoms of the common cores of 32 pairs of homologous proteins (see Table II).

Immunoglobulin CDRs

Antibody specificity is determined by 6 loops (CDRs)

Billions of different ligands

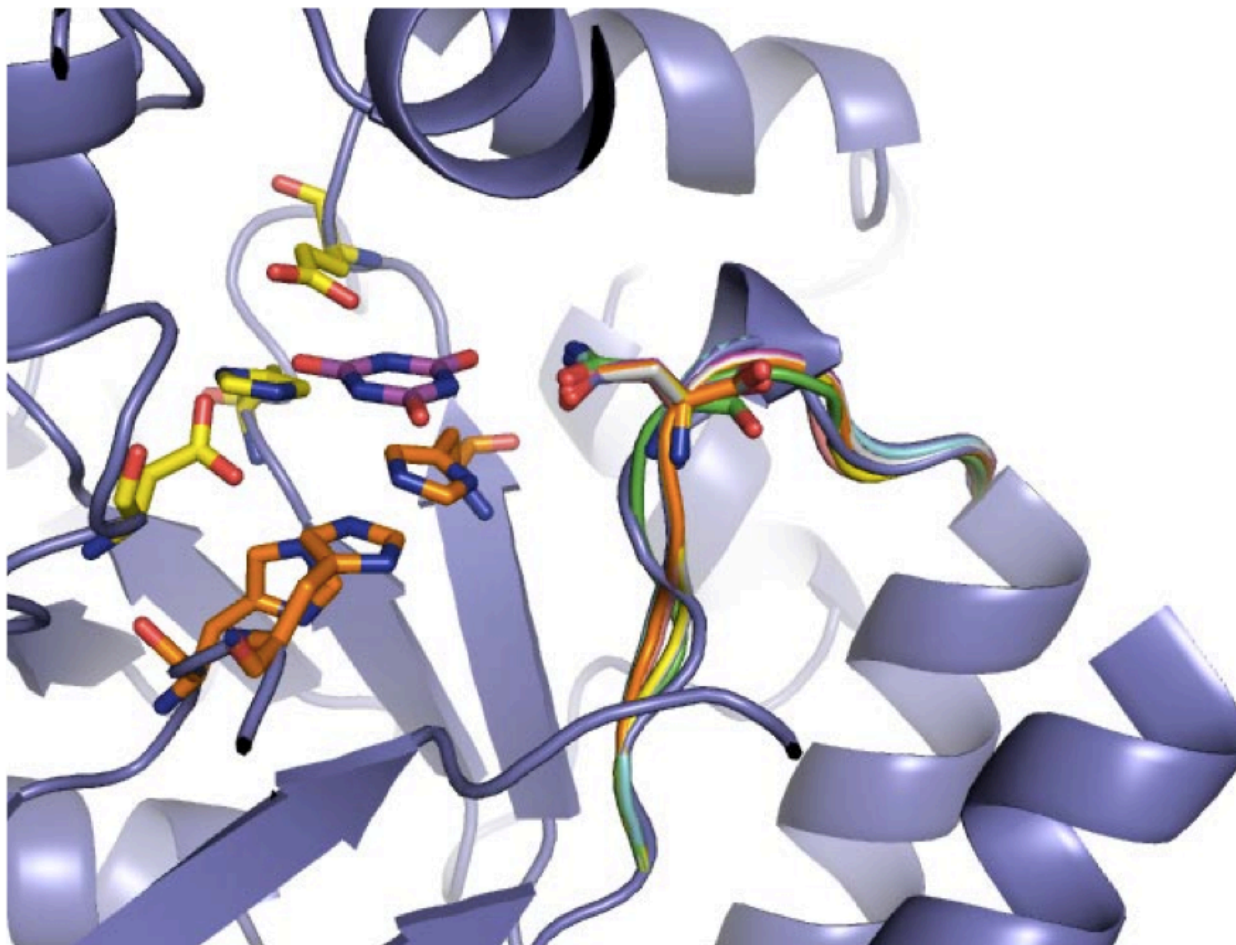
Ab-specific rules for bb conformation



Enzymes

Rosetta

Enzyme function and
regulation

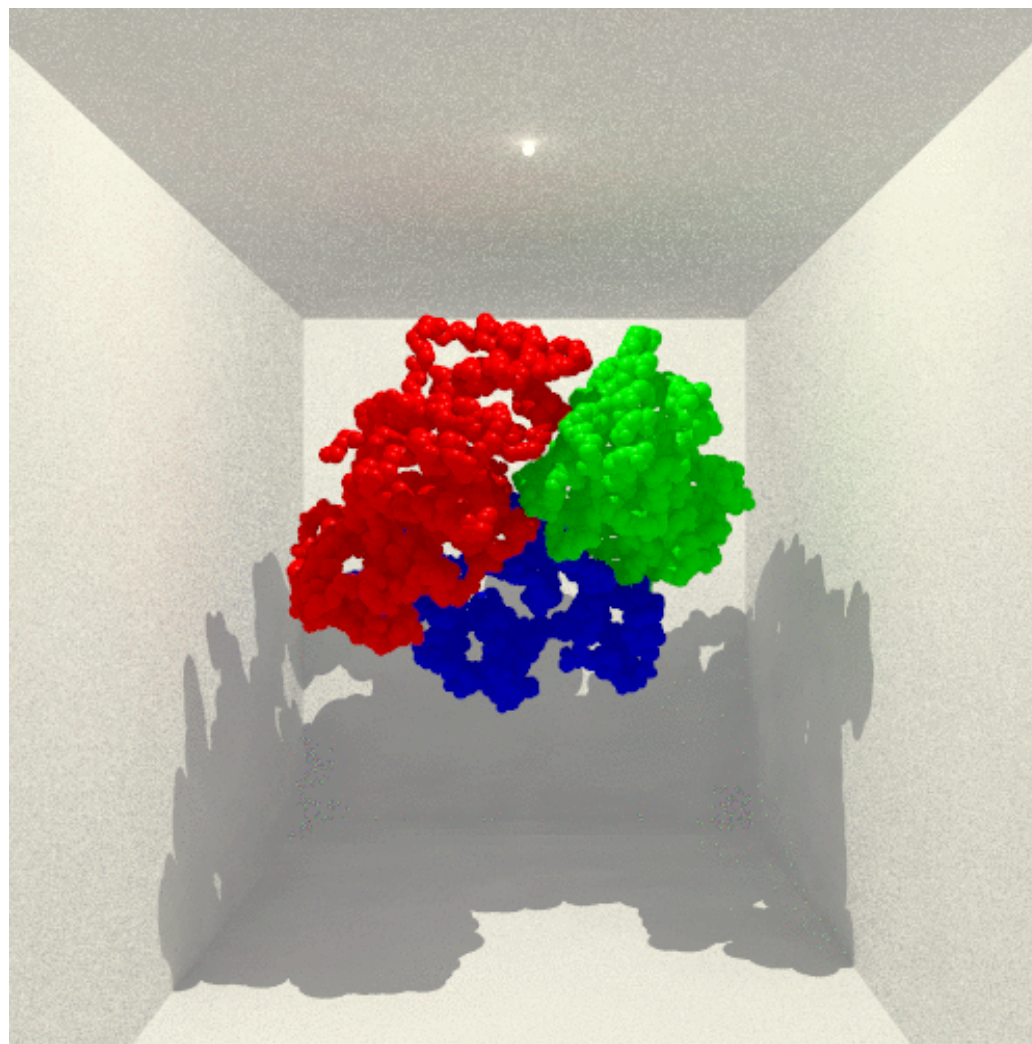


Alteration of enzyme specificity by computational loop remodeling and design

Paul M. Murphy^{a,b,c}, Jill M. Bolduc^d, Jasmine L. Gallaher^e, Barry L. Stoddard^d, and David Baker^{a,e,1}

Docking

Clashes and bad geometries
must be fixed before docking!



Modeller

protocols for loop building and refinement

Basic idea:

generate hundreds of loops

pick one based on energy **AND** cluster size

Neither fast nor accurate

Fread

Loop database

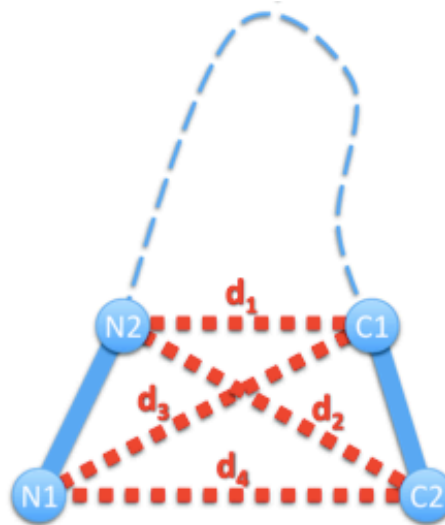
Sequence and
anchor point similarity

<http://opig.stats.ox.ac.uk/webapps/fread/php/>

Fread

Loop database

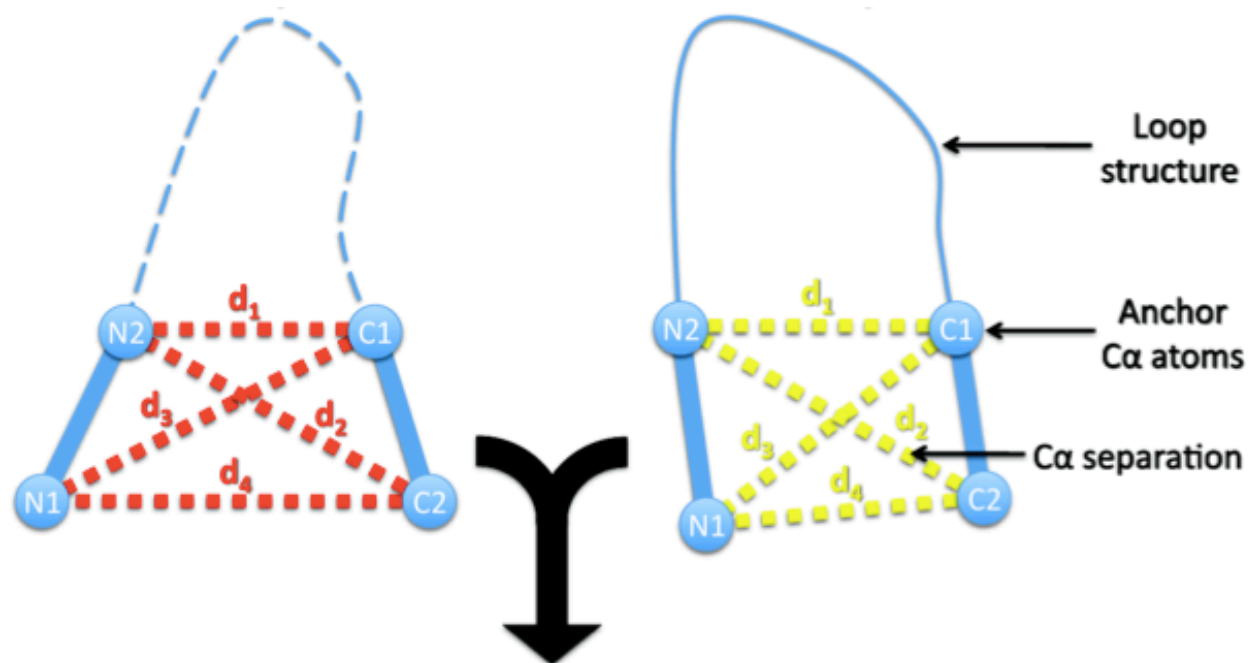
Sequence and
anchor point similarity



Fread

Loop database

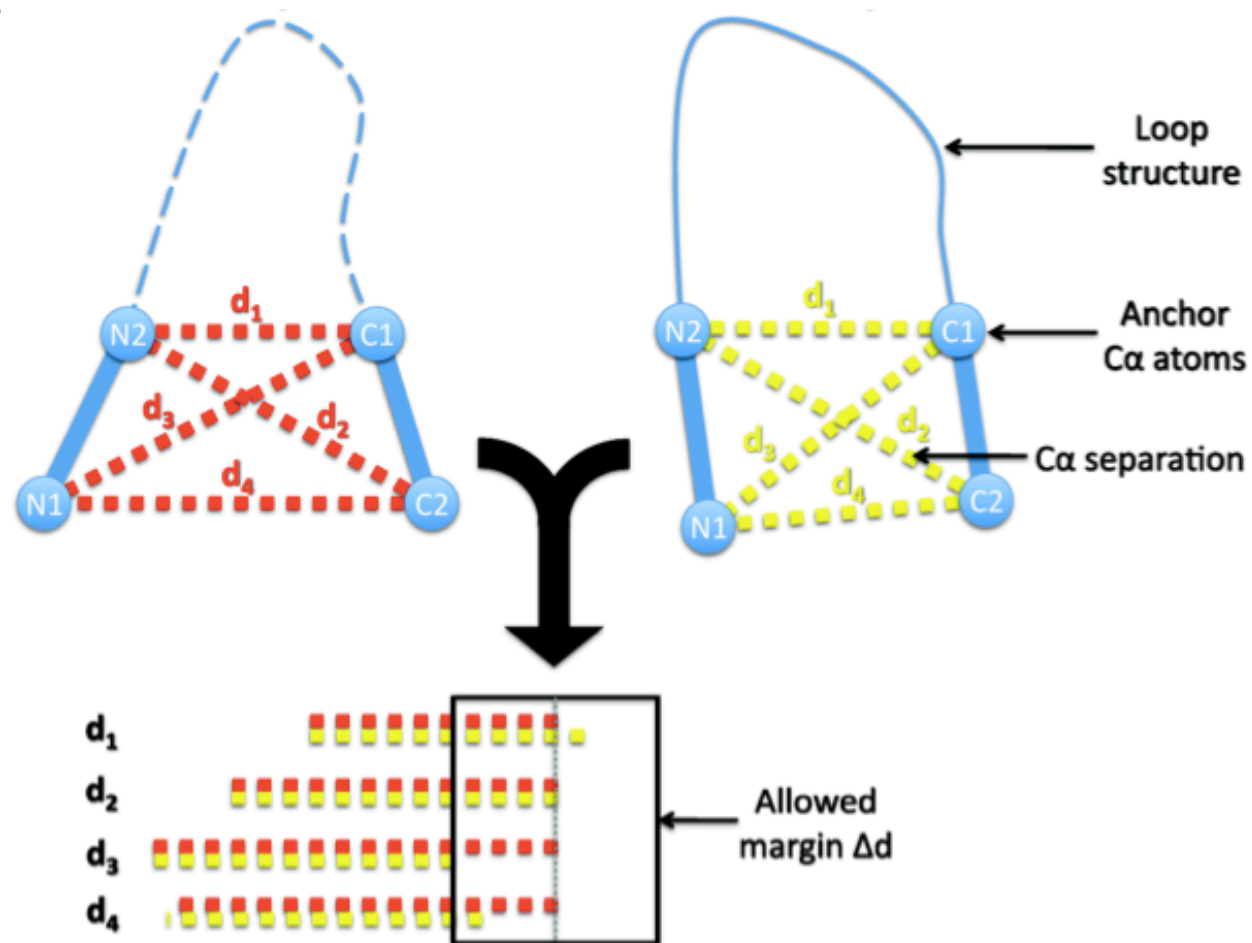
Sequence and
anchor point similarity



Fread

Loop database

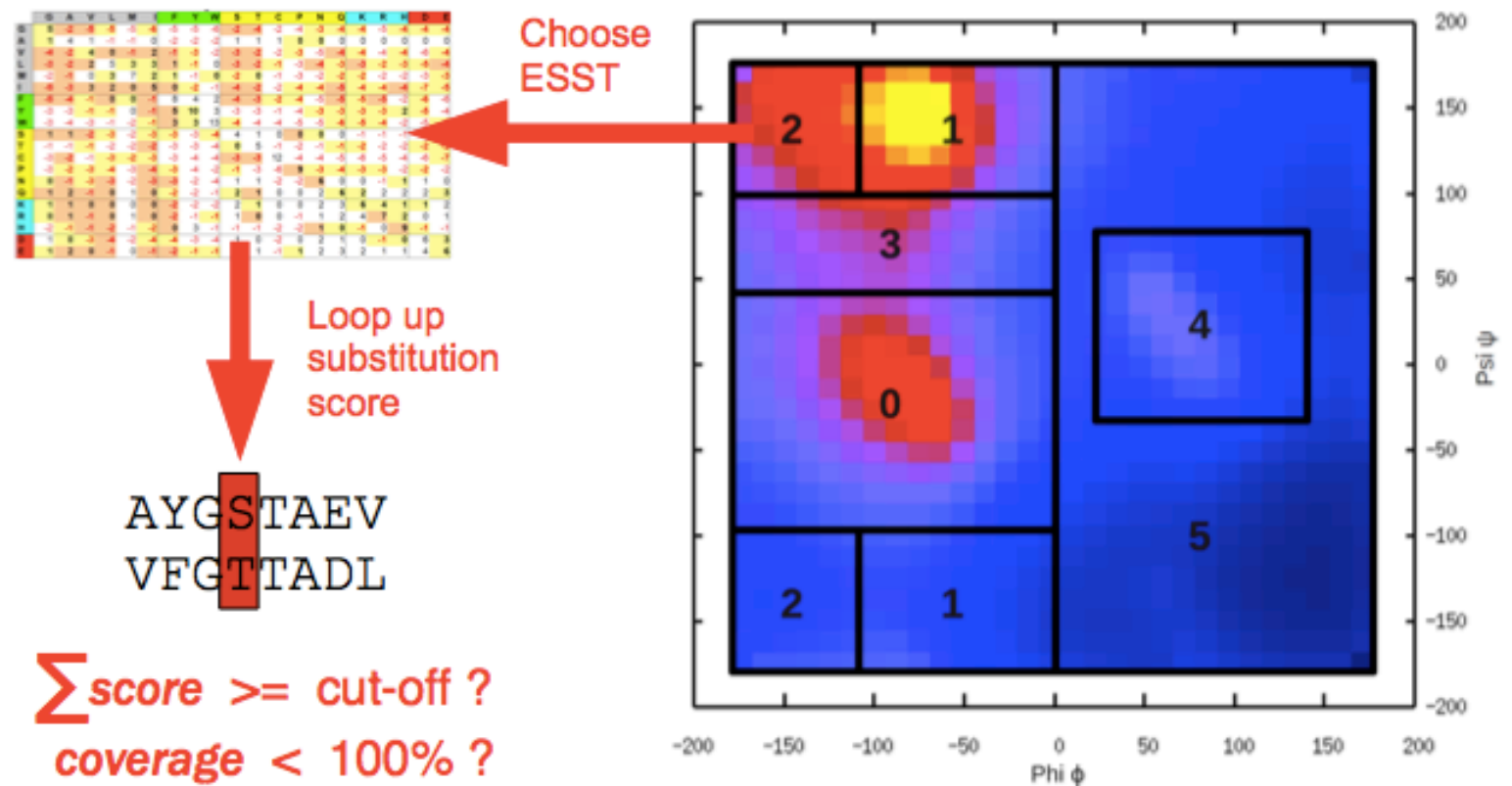
Sequence and
anchor point similarity



Fread

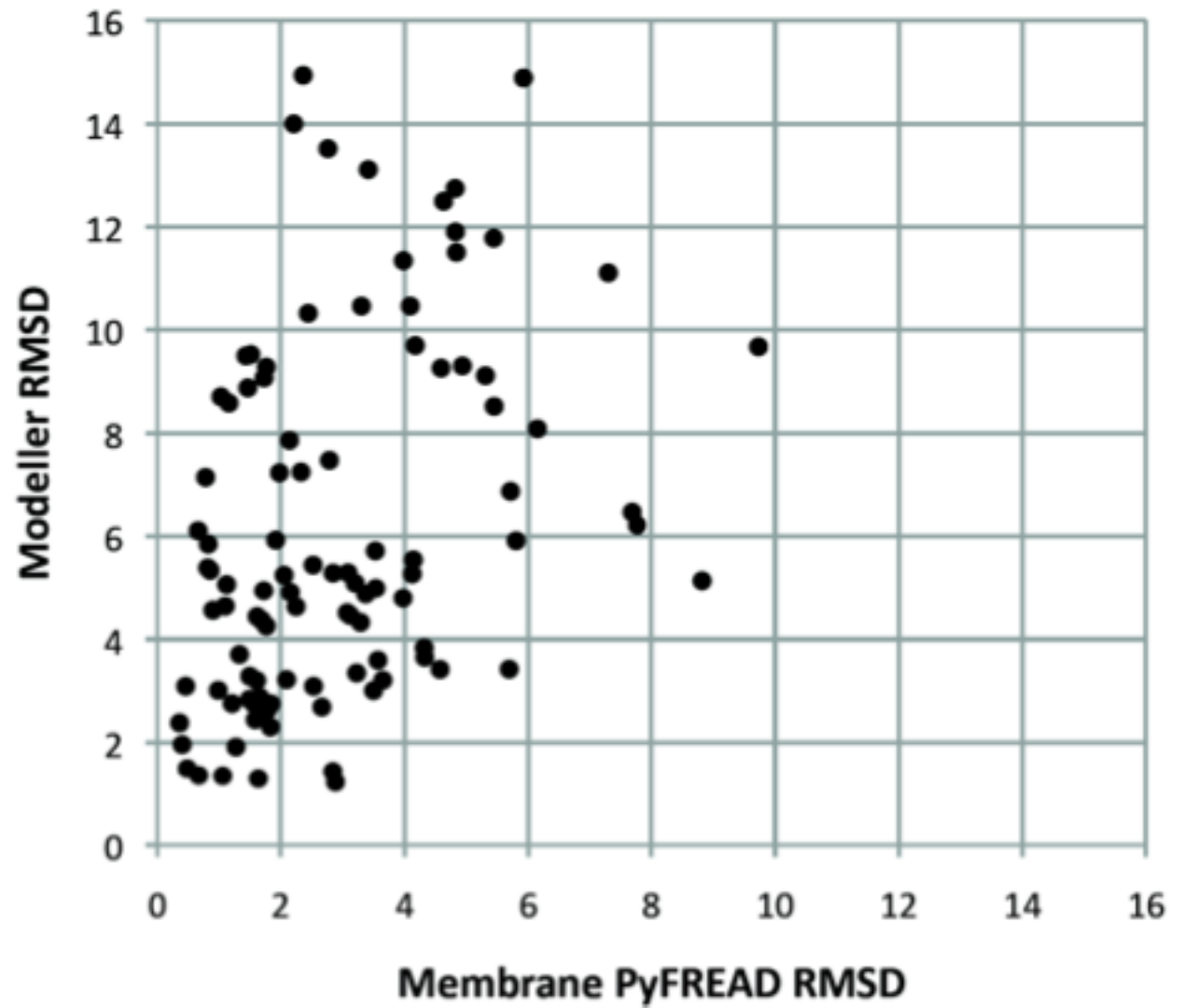
Loop database

Sequence and
anchor point similarity



Fread

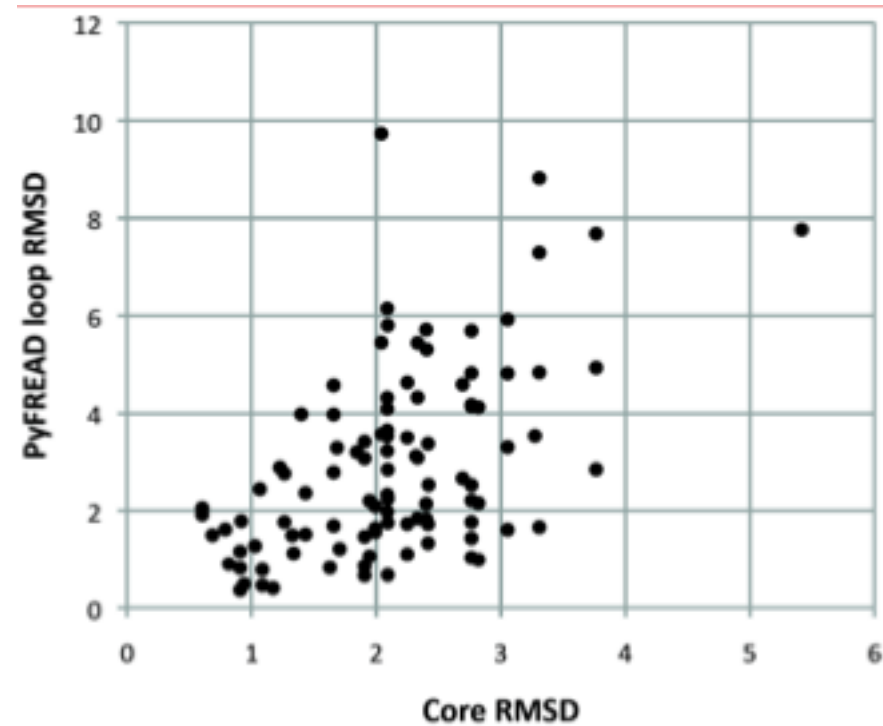
Outperforms modeller



Fread

Outperforms modeller

Anchor RMSD is important

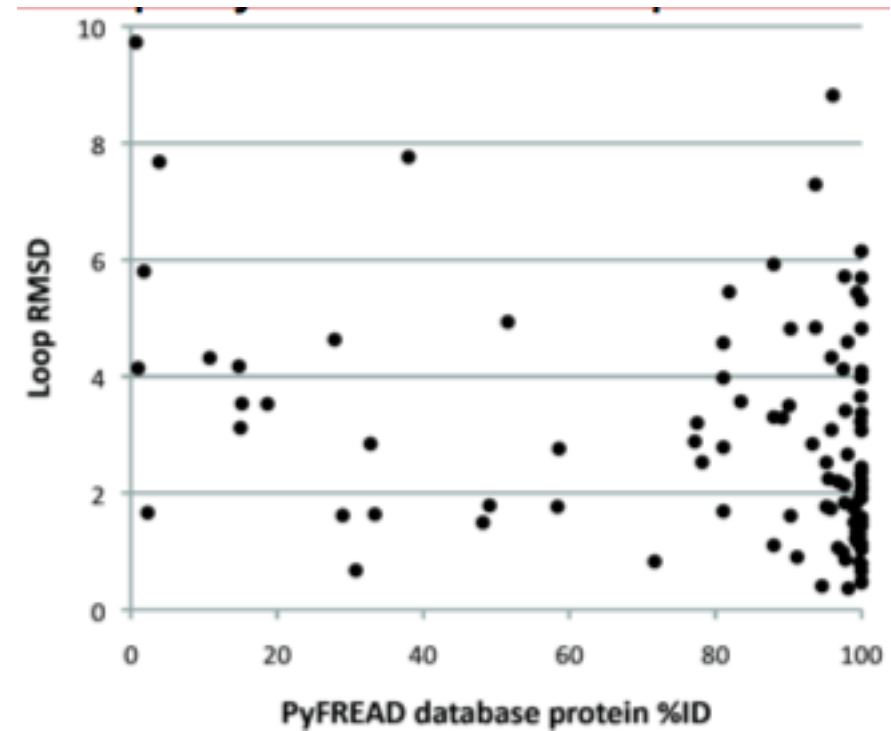


Fread

Outperforms modeller

Anchor RMSD is important

Identity has small influence



Fread

Outperforms modeller

Anchor RMSD is important

Identity has small influence

no template or anchor RMSD too high?

KIC

Kinematic Closure

De Novo

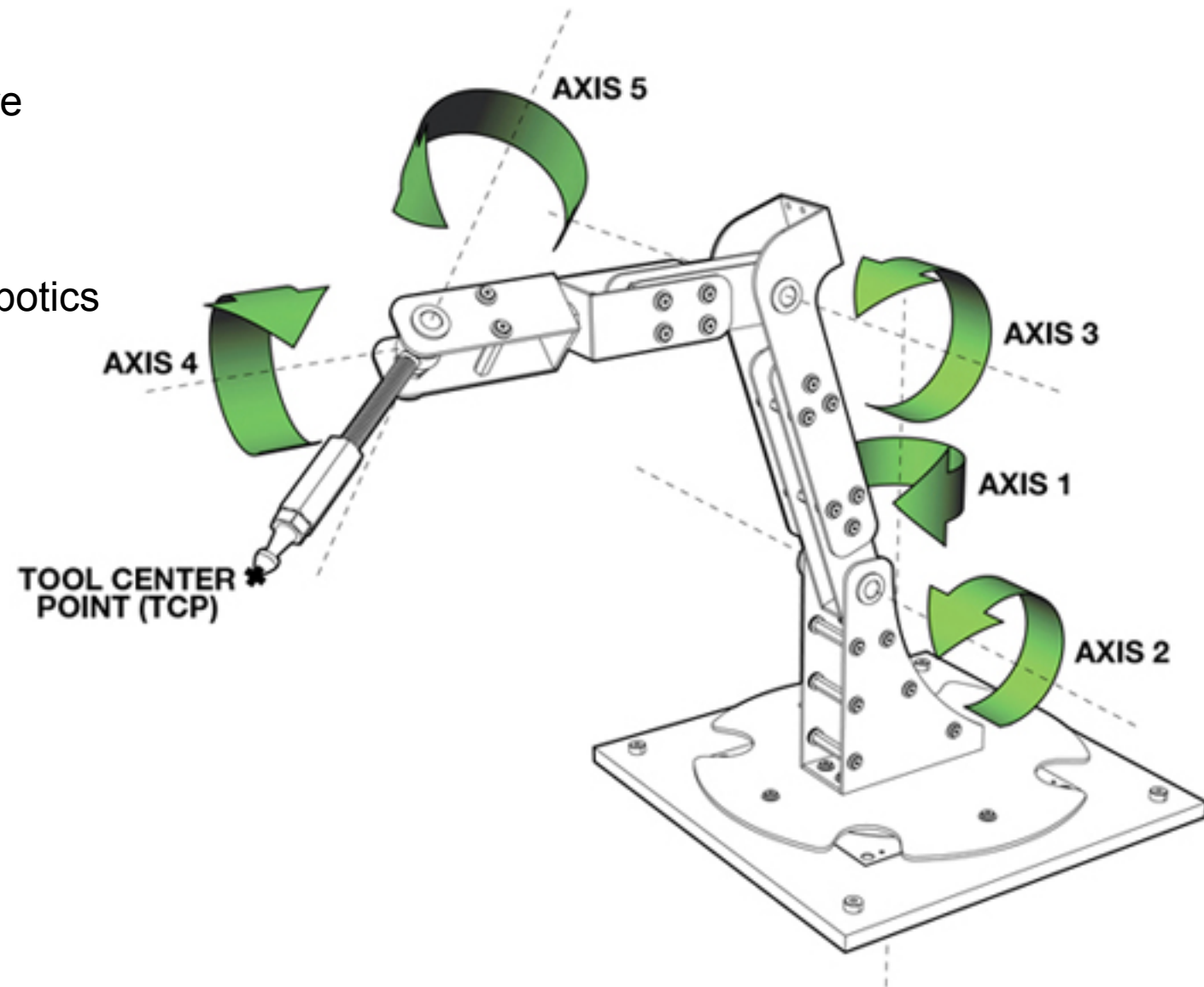
Inspired from Robotics

KIC

Kinematic Closure

De Novo

Inspired from Robotics



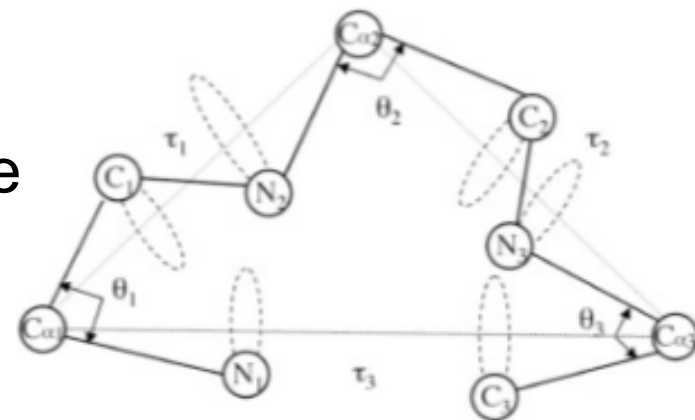
KIC

From robotics: Analytical solution of loop closure for **6** degrees of freedom

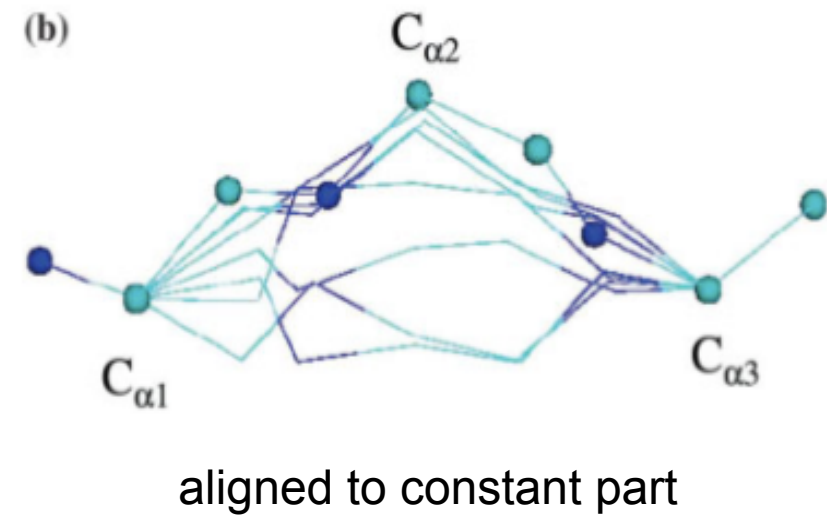
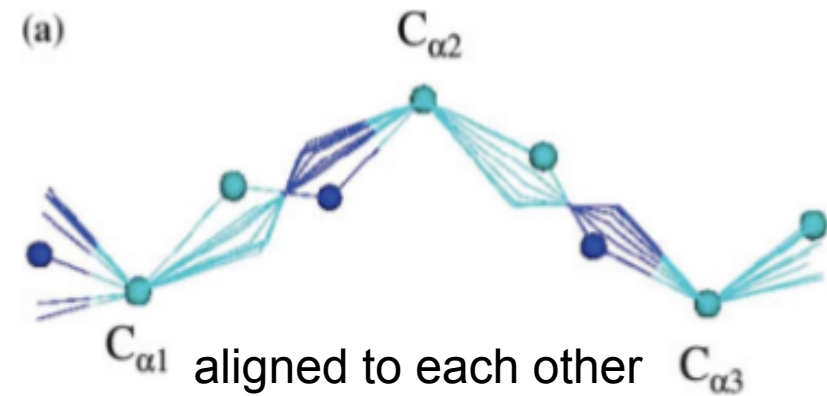
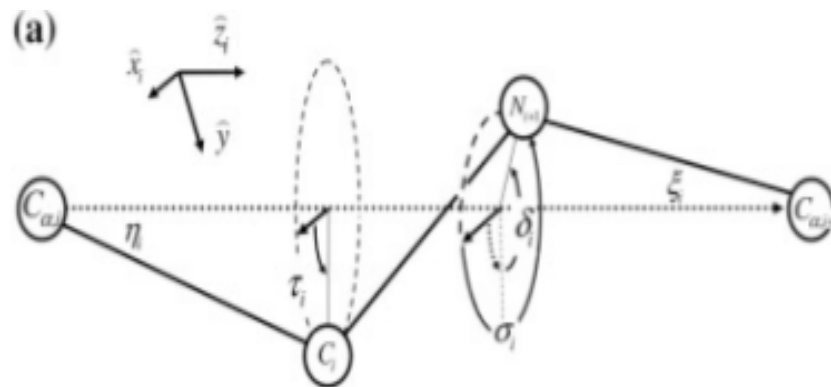
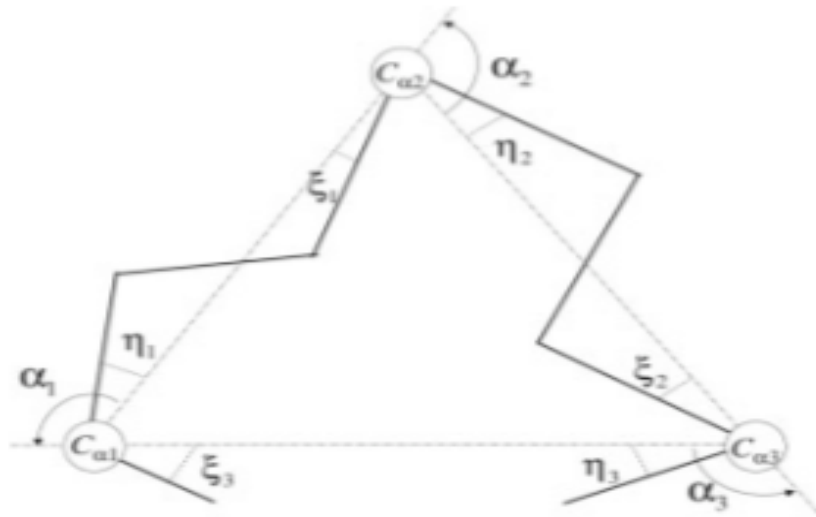
Challenge:

- find **analytical formulation** to extract
- **all possible** backbone structures of a chain segment, that are
- **geometrically consistent** with preceding and following parts of the given structure.

Setup:



Solutions

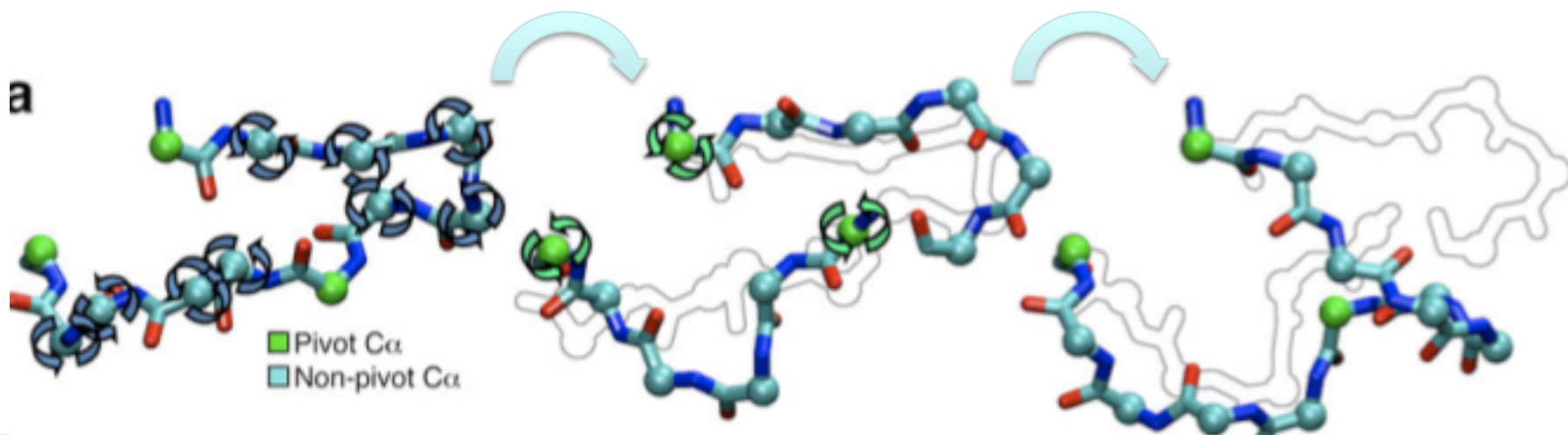


KIC

- Analytical solution of loop closure for **6** degrees of freedom
- Extension:** analytical determination of all mechanically accessible conformations for 6 torsions of a peptide chain of **any length** (e.g. 25 residues)

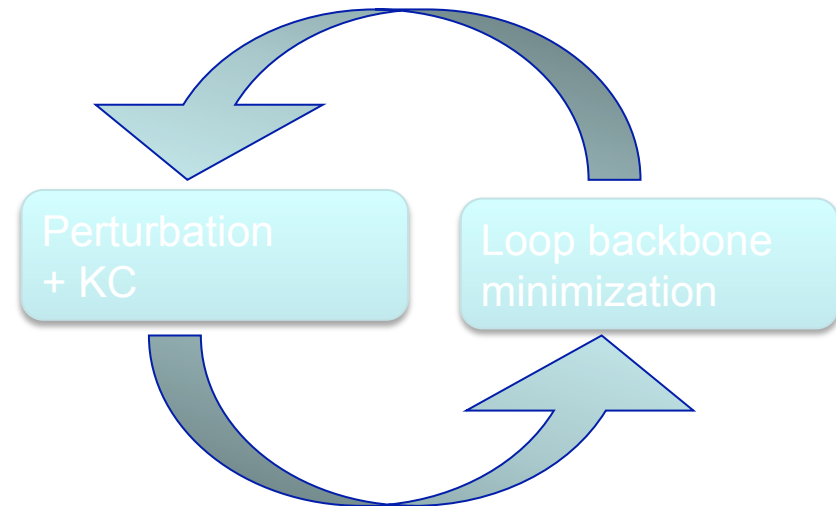
(1) Randomly perturb non-pivot positions

(2) Apply KC to pivot positions



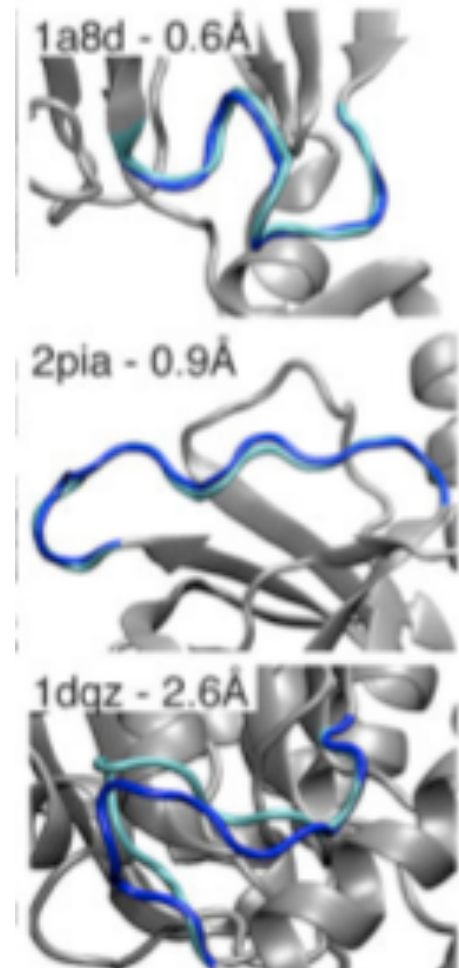
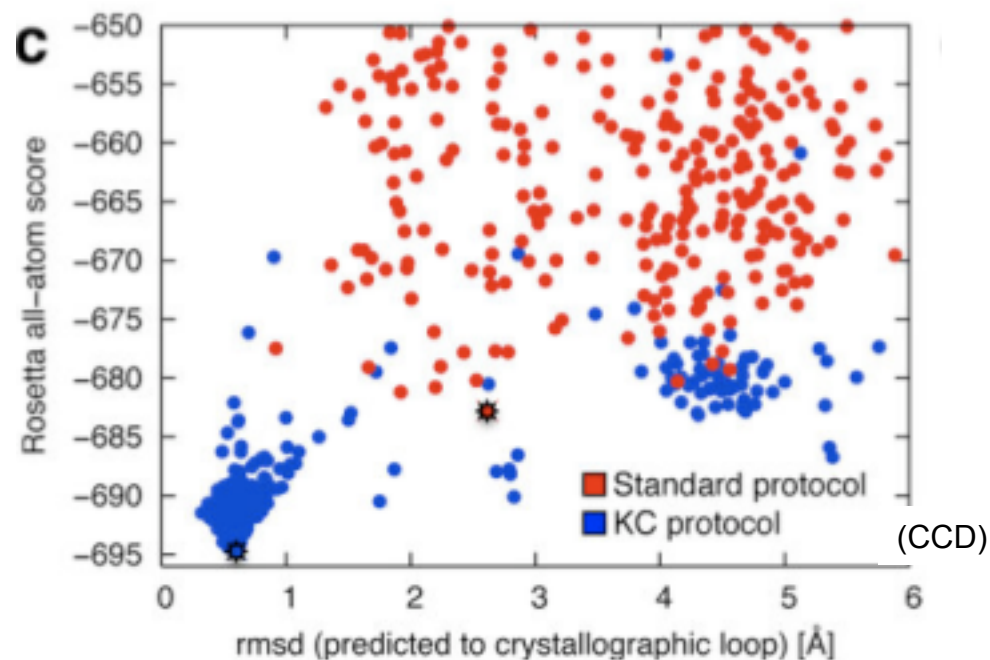
KIC in Rosetta

- Embedded into MCM protocol (low-res + high-res)
 - 720 steps
 - Repeat 1000 times



KIC accuracy

- Improves median modeling quality from 2.0Å to **0.8Å** RMSD (on set of 25 loops)



Protein refinement

Use for MD, docking

Improve appearance

Check for non-trivial clashes

Overall accuracy not affected

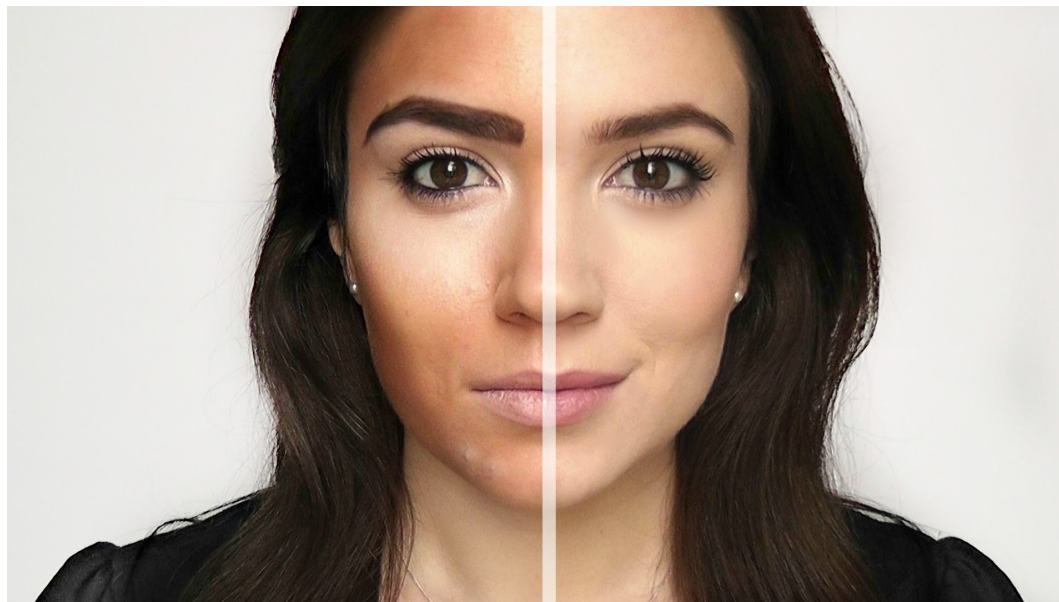
Protein refinement

Use for MD, docking

Improve appearance

Check for non-trivial clashes

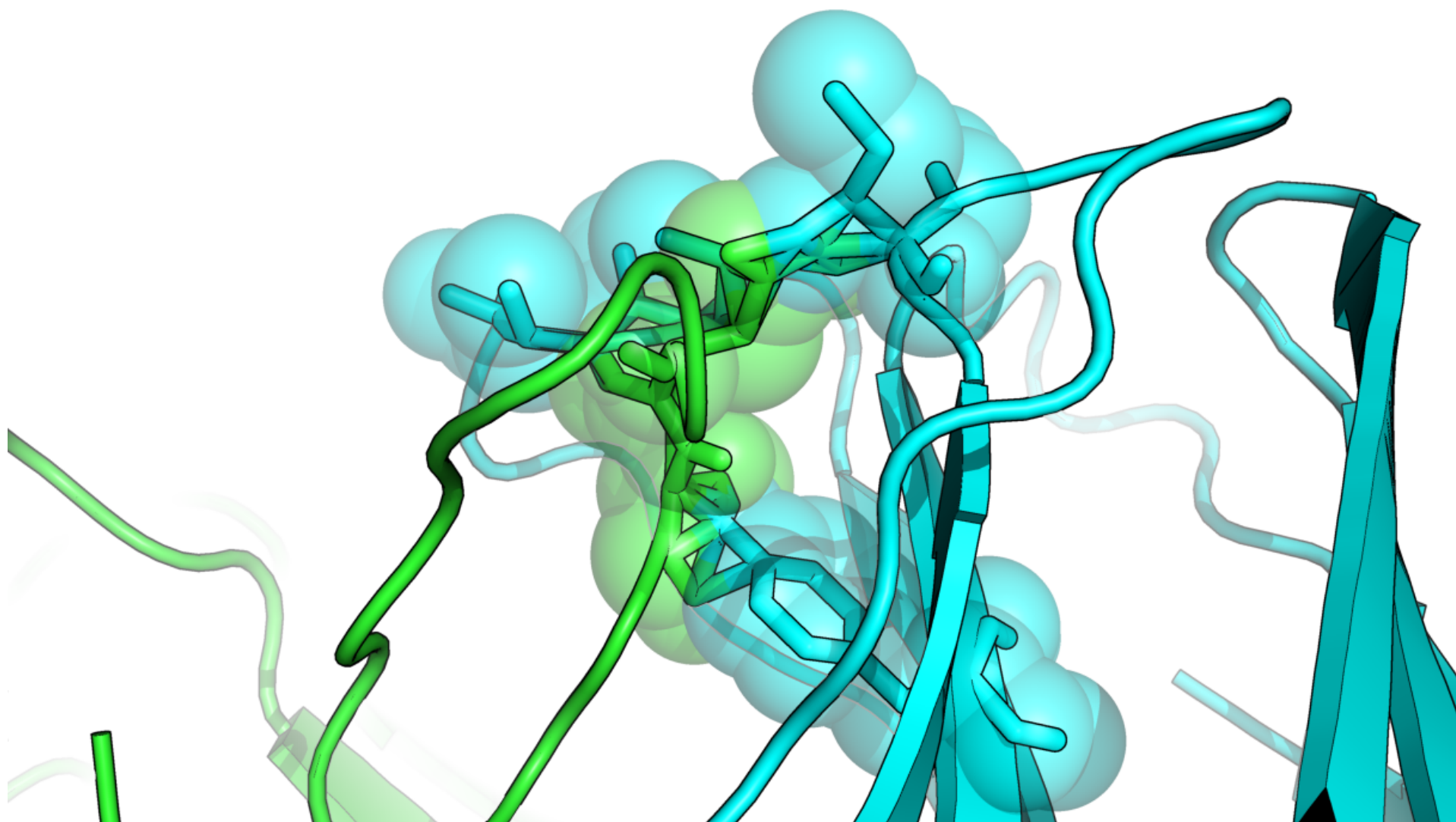
Overall accuracy not affected



Like makeup, it's there but should look natural

refinement example

CENTERFO
R BIOLOGI
CAL SEQU
ENCE ANA
LYSIS CBS



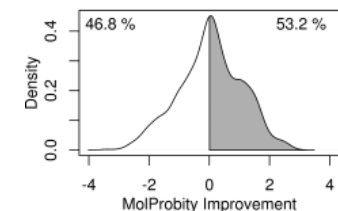
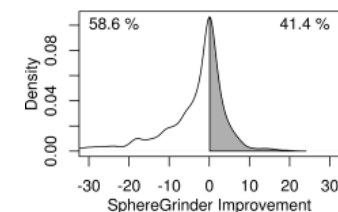
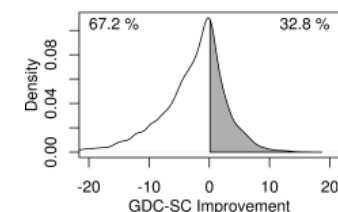
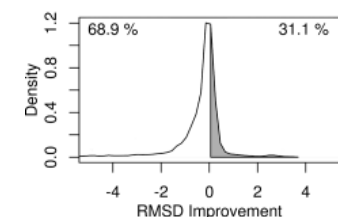
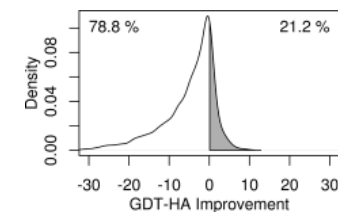
refinement example

CENTERFO
R BIOLOGI
CAL SEQU
ENCE ANA
LYSIS CBS



refinement results in CASP

"distributions skewed to the left, indicating there are more big failures than big improvements."



Evaluation of predictions in the CASP10 model refinement category

Timothy Nugent, Domenico Cozzetto, and David T. Jones*

Resources for Loop Modeling

Libraries

Swiss PDB viewer / Swiss modeller (<http://spdbv.vital-it.ch/>)

Fread (<http://opig.stats.ox.ac.uk/webapps/fread/php/>)

Brix (<http://brix.crg.es/>)

Ab Initio

Rosetta (<https://www.rosettacommons.org>)

Modeller (<https://salilab.org/modeller/>)

Resources for refining

Swiss PDB viewer / Swiss modeller (<http://spdbv.vital-it.ch/>)

refine (<http://www.cbs.dtu.dk/biotools/Refine/>)

KobaMin (<http://csb.stanford.edu/kobamin/>)

Modeller (<https://salilab.org/modeller/>)

refine (<http://www.cbs.dtu.dk/biotools/Refine/>)

ModRefiner (<http://zhanglab.ccmb.med.umich.edu/ModRefiner/>)

Summary

- ✓ Re-model the **important loops**
 - ✓ Use **loop libraries** for short loops
 - ✓ If no template or loop too long, use **loop closure**
 - ✓ **Refinement** is often needed but
 - ✓ **No improvement** expected
 - ✓ Refine as little as you can!
-